

WHAT IS CLAIMED IS:

1. A method for manipulation, storage, modeling, visualization and quantification of datasets comprising:
- 5 providing a plurality of target strings comprising datasets;
- generating a comparison string comprising a dataset using an iterative algorithm, such that the comparison string is calculated from a point in any set of points that can serve as the domain of an iterative function;
- scoring of the comparison string by evaluating a function having the
- 10 comparison string and one of the plurality of target strings as inputs, such that the evaluation may be repeated for a number of the other plurality of target strings;
- mapping or marking the point if the score or some other property corresponding to the point meets some relevant criteria;
- repeating the generating, scoring, and mapping or marking for a plurality of
- 15 comparison strings if desired; and
- examining a subregion with higher resolution if points in the subregion are of interest.
2. The method of claim 1, wherein the step of providing the comparison
- 20 string comprises transforming the numbers of the comparison string to have values within a set of interest.
3. The method of claim 1, wherein the set of points comprises a region of the complex plane.
- 25
4. The method of claim 3, wherein the set of points further comprises points in and/or near the Mandelbrot Set or a Julia Set.
5. The method of claim 1, wherein the step of generating the comparison
- 30 string comprises laying a grid over the set of points.
6. The method of claim 1, wherein the step of generating the comparison string comprises restarting the step of generating the comparison string if the iteration

has become unbounded.

7. The method of claim 1, wherein the step of generating the comparison string comprises generating a comparison string of any length.

5

8. The method of claim 1, wherein the step of scoring comprises preliminary testing of properties of the comparison string alone as criteria to initiate scoring.

10 9. The method of claim 1, wherein the step of scoring comprises some test of the comparison string using the target string.

10. The method of claim 9, wherein not all of the numbers in the comparison string or the target string must be used in the test.

15

11. The method of claim 1, wherein the step of scoring comprises a one-to-one comparison between corresponding numbers in the target string and the comparison string.

20 12. The method of claim 11, wherein the one-to-one comparison may be between corresponding sequential or non-sequential numbers in the target string and the comparison string.

25 13. The method of claim 1, wherein the step of scoring involves studying the behavior of the scoring function, such as determining the function's minima and maxima.

14. The method of claim 13, wherein only the comparison string is used as relevant input to the scoring function.

30

15. The method of claim 1, wherein the step of mapping or marking comprises storing the coordinates of the point corresponding to the target string or properties of the comparison string in memory, a database or a table.

16. The method of claim 1, wherein the step of mapping or marking comprises marking the point on a visual display by changing some graphical property of the corresponding pixel, such as color.

5 17. The method of claim 1, wherein the criteria comprises the comparison string having the highest score, where the score is based on some similarity measure to the target string.

10 18. The method of claim 1, wherein the step of examining the subregion comprises changing the format of the target and/or comparison string in order to improve the precision and resolution of the method.

15 19. The method of claim 18, wherein the methodology used in the reformatting process is based on methodologies such as Simulated Annealing, Hill Climbing Algorithms, Genetic Algorithms, or Evolutionary Programming Methods.

20. The method of claim 19, wherein the reformatting process is automated.

20 21. The method of claim 1, wherein the points of interest are analyzed and/or compared by examining, either visually or mathematically, their relative locations and/or absolute locations within the region.

25 22. The method of claim 1, wherein the points of interest are analyzed and/or compared by examining, either visually or mathematically, metrics other than location.

23. The method of claim 22, wherein the metrics can be represented by graphic properties such as shading.

30

24. The method of claim 1, wherein the step of examining a subregion further comprises repeating the examining step for smaller subregions.

25. The method of claim 1, wherein the uses for the method comprise analyzing large datasets, such as for DNA sequence data, protein sequence data, gene expression datasets, demographic data, statistical data, and clinical (patient) data.

5 26. The method of claim 1, wherein the uses of the method comprise analyzing datasets consisting of heterogeneous data, such as both gene expression data and clinical (patient) data.

27. The method of claim 1, wherein the uses for the method comprise data
10 compression.

28. The method of claim 1, wherein the steps may be automated.

29. The method of claim 1, wherein separates processes involved in
15 generating each comparison string, scoring each comparison string, or transforming each comparison string or target string to a value set of interest may be processed simultaneously by a plurality of processors.

30. A method for manipulation, storage, modeling, visualization and
20 quantification of datasets comprising:
providing a plurality of target strings comprising datasets;
generating a comparison string comprising a dataset using an iterative algorithm, such that the comparison string is calculated from a point in a region of the complex plane and the numbers of the comparison string are transformed to have
25 values within a set of interest;

scoring of the comparison string by evaluating a function having the comparison string and one of the plurality of target strings as inputs, such that the evaluation may be repeated for a number of the other plurality of target strings;

mapping or marking the point if the score or some other property
30 corresponding to the point meets some relevant criteria, such that the coordinates of the point corresponding to the target string or properties of the comparison string are stored in memory, a database or a table, or the point is marked on a visual display by

changing some graphical property of the corresponding pixel, and wherein the relevant criteria comprises the comparison string having the highest score, where the score is based on some similarity measure to the target string;

- repeating the generating, scoring, and mapping or marking for a plurality of
- 5 comparison strings if desired; and
- examining a subregion with higher resolution if points in the subregion are of interest, wherein the points of interest are analyzed and/or compared by examining, either visually or mathematically, their relative locations and/or absolute locations within the region or other metrics representing the graphic properties of the
- 10 corresponding comparison strings.

Approved for Release